



HAL
open science

Event-driven ECG classification using an open-source, LC-ADC based non-uniformly sampled dataset

M. Saeed, Q. Wang, O. Märtens, Benoit Larras, Antoine Frappé, B. Cardiff,
D. John

► To cite this version:

M. Saeed, Q. Wang, O. Märtens, Benoit Larras, Antoine Frappé, et al.. Event-driven ECG classification using an open-source, LC-ADC based non-uniformly sampled dataset. 53rd IEEE International Symposium on Circuits and Systems, ISCAS 2021, May 2021, Daegu, South Korea. pp.9401333, 10.1109/ISCAS51556.2021.9401333 . hal-03362265

HAL Id: hal-03362265

<https://hal.science/hal-03362265v1>

Submitted on 17 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Event-Driven ECG Classification Using an Open-Source, LC-ADC Based Non-Uniformly Sampled Dataset

Maryam Saeed*, Qingyuan Wang*, Olev Märtens[†], Benoit Larras[‡], Antoine Frappé[‡], Barry Cardiff[§], Deepu John[§]
*[§]University College Dublin, Ireland, [†]Tallinn University of Technology, Estonia, [‡]Université de Lille, France

Abstract—In this article, non-uniformly sampled electrocardiogram (ECG) signals obtained from level-crossing analog-to-digital converters (LC-ADCs) are analyzed for event-driven classification and compression performance. The signal compression results show that it is important to assess the distortion in event-driven signals when simulating LC-ADC models, especially at lower resolutions and larger quantization steps. The effects of varying the LC-ADC parameters for the application of cardiac arrhythmia classifiers are also assessed using an artificial neural network (ANN) and the MIT-BIH Arrhythmia Database. In comparison with uniformly-sampled data, it is possible to achieve comparable classification accuracy at a much lower complexity with event-driven ECG signals. The results show the best event-driven model achieves over 97% accuracy with 79% reduction in ANN complexity with signal-to-distortion ratio (S/D) ≥ 21 dB. For S/D < 21 dB, the best event-driven model achieves 93% accuracy with a 96% reduction in ANN complexity. An open-source event-driven arrhythmia database is also presented.

Index Terms—LC-ADC, cardiac arrhythmia classification, artificial neural networks, wearable sensors, event-driven data

I. INTRODUCTION

Long-term health monitoring using wireless biomedical devices has been enabled by the evolution in the Internet of Things (IoT) and is becoming increasingly ubiquitous in recent times [1]. Studies have shown that the energy costs to wirelessly transmit data in a standard biosensor is “orders of magnitude greater than any other function” [2]. To prolong the battery life of wireless sensors, feature extraction, compression techniques are typically incorporated on-chip after data conversion using an analog-to-digital converter (ADC) [3], [4]. In recent years, a new class of ADCs, called level-crossing ADCs (LC-ADCs), has been developed which embed compression into the data acquisition stage. They generate non-uniformly sampled data that can reproduce the original data, from inherently sparse signals, such as electrocardiogram (ECG) signals. LC-ADCs sample data based on the slope of the signal, i.e., they only sample when the rate of change of signal is above a pre-defined level, indicating a level-cross. This keeps the ADC from sampling across low-activity regions of the signal, thus reducing the data rate and saving power. LC-ADCs have the advantage of built-in compression and lower processing costs in comparison with Nyquist sampling ADCs.

This work is supported by 1) JEDAI project under the Chist-Era Program; 2) Schlumberger Foundation’s Faculty for the Future Program and 3) Irish Research Council under the New Foundations Scheme

A recent study compared the power efficiency of an LC-ADC with a traditional SAR ADC using various biosignals [5], which showed that LC-ADCs are more power-efficient than standard ADCs for low-to-medium resolution applications.

Many low-power LC-ADC designs have been proposed in recent years for various biomedical applications [6] [7] [8]. LC-ADCs have also been used in low-power QRS-detection of ECG signals [9] [10]. More recently, clinically acceptable arrhythmia classification has been achieved using event-driven data from LC-ADCs [11]. To the best of our knowledge, performance analysis of various LC-ADC designs for event-driven ECG classification has not yet been done. This article aims to present a comparative analysis of several LC-ADC designs in terms of compression ratio, percentage root-mean-square difference, signal-to-distortion ratio, and their performance on an event-driven cardiac arrhythmia classifier. Finally, we intend to provide an insight into the best accuracy-complexity trade-off for this application. The major contributions of this article are as follows: a) analysis of three level-crossing ADC models in terms of compression parameters to assess the limits of clinically acceptable event-driven ECG signal b) proposed an ANN model for classification of cardiac arrhythmia from event-driven ECG signal c) an open-source event-driven ECG database derived from MIT-BIH database using existing LC-ADC models with arrhythmia annotations.

The rest of the article is organized as follows: Section II presents the compression performance of three LC-ADC designs. Section III assesses in detail the performance of a cardiac arrhythmia classifier using event-driven data from these LC-ADCs designs. It also presents an event-driven arrhythmia database. Finally, conclusions are drawn in Section IV.

II. COMPRESSION ANALYSIS OF LC-ADC DESIGNS FOR ECG APPLICATIONS

A. Level-Crossing Sampling

Fig. 1 presents the basic architecture of an LC-ADC and the analog ECG signal with the event-driven output. The ECG signal range is divided into a fixed number of quantization levels. The LC-ADC takes a sample based on two in-process quantization levels: upper quantization level (U_QL) and lower quantization level (L_QL). The distance between these

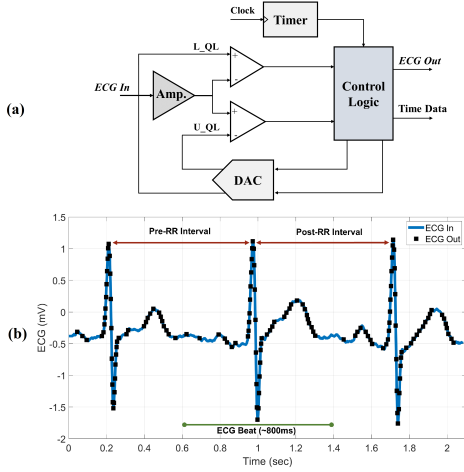


Fig. 1. (a) LC-ADC architecture (b) ECG analog signal, the LC-ADC output and the RR intervals.

quantization levels is defined by the LSB, also known as the quantization step:

$$LSB = \frac{A_{FS}}{2^M} \quad (1)$$

where A_{FS} is the full-scale voltage range of the ECG signal, M is the resolution of the LC-ADC and 2^M is the total number of quantization levels. In some LC-ADC designs, the full-scale range A_{FS} is often multiplied by a factor of 2. The input analog signal must always be surrounded by the two in-process quantization levels in an LC-ADC. Whenever the signal crosses U_QL , a sample is taken and both in-process quantization levels are increased by one LSB. Similarly, whenever L_QL is crossed, both in-process quantization levels are decreased by one LSB. The control logic block in Fig. 1 manages this update. This way, an LC-ADC creates continuous-in-time discrete-in-amplitude ECG signals. To monitor the time between two LC-ADC samples, a local timer is used, which passes the time information at each sample and then resets. Often, the distance between the in-process quantization levels can be multiplied by a factor k , which is defined as:

$$k = ku * LSB \quad (2)$$

where ku is an integer greater than or equal to 1. For higher resolution LC-ADCs, k is often chosen as $2LSB$ or greater.

B. LC-ADC Models

Three different LC-ADC design models have been chosen for the compression analysis. The models will be assessed using the MIT-BIH Arrhythmia dataset [12], which has been sampled at 360Hz with an input voltage range of [-5mV, 5mV] and an 11-bit ADC resolution. The LC-ADC models are defined as follows:

- 1) *MDL1*: The LC-ADC model 1, hereby named MDL1, has been taken from the LC-ADC design in [11]. In this model, the MIT-BIH Arrhythmia database is converted to event-driven ECG using the LSB definition in (1)

multiplied by a factor of 2. The ECG data is not up-sampled before the event-driven sampling.

- 2) *MDL2*: The LC-ADC model 2, hereby named MDL2, is designed as in [10]. In this model, the database is first up-sampled to 5.67kHz and uses the LSB definition in (1).
- 3) *MDL3*: The LC-ADC model 3, hereby named MDL3, has been designed as in [8]. In this model, the database is first up-sampled to 10kHz and uses the LSB definition in (1).

C. Performance Metrics

To measure the signal distortion in the LC-ADC output, its output samples are interpolated using linear interpolation to reconstruct the ECG signal, as in [8]. The three performance parameters are as follows:

- 1) *Compression Ratio (CR)*: The compression ratio provides information about the degree of compression between the compressed and uncompressed signal. It is defined as:

$$CR = \frac{m}{n} \quad (3)$$

where m and n are the number of bits in the original and compressed signal, respectively.

- 2) *Percentage Root Mean Squared Difference (PRD)*: The PRD measure evaluates the distortion between the original signal and the reconstructed signal. The normalized PRD is defined as:

$$PRD = \sqrt{\frac{\sum_{n=1}^N (x[n] - \hat{x}[n])^2}{\sum_{n=1}^N (x[n] - \bar{x}[n])^2}} * 100 \quad (4)$$

where $x[n]$, $\hat{x}[n]$, and $\bar{x}[n]$ are the original signal, reconstructed signal and the mean of the signal $x[n]$, respectively. N is the total number of samples that model the input analog signal.

- 3) *Signal-to-Distortion Ratio (S/D)*: Another metric to measure distortion is the temporal signal-to-distortion ratio defined as:

$$S/D = 10 \log_{10} \left(\frac{\sum_{n=1}^N (x[n] - \bar{x}[n])^2}{\sum_{n=1}^N (x[n] - \hat{x}[n])^2} \right) \quad (5)$$

In [13], the clinically acceptable values for PRD and S/D in ECG signal reconstruction are also reported. "Good" quality signals must follow the following criteria:

$$PRD \leq 9\% \text{ or } S/D \geq 21dB \quad (6)$$

D. Compression Performance of LC-ADC Models

The three LC-ADC models are designed and simulated using MATLAB. To match the MIT-BIH Arrhythmia database voltage range, A_{FS} in (1) is chosen as 10mV. The resolution of the LC-ADC, M , is varied from 7 bits to 11 bits and k in (2) is varied from 1LSB to 4LSB. Table I shows the average compression results for all models calculated over

TABLE I
COMPRESSION PERFORMANCE OF LC-ADC MODELS

LC-ADC Model	M	k	CR	PRD(%)	S/D(dB)
MDL1	11	1	1.574	1.886	35.029
		2	2.722	3.595	29.471
		3	4.026	5.255	26.179
		4	5.277	6.901	23.809
	10	1	2.722	3.595	29.471
		2	5.277	6.901	23.809
		3	7.382	10.387	20.263
		4	9.269	14.221	17.535
	9	1	5.277	6.901	23.809
		2	9.269	14.221	17.535
		3	13.070	22.799	13.432
		4	17.050	32.146	10.496
8	1	9.269	14.221	17.535	
	2	17.050	32.146	10.496	
	3	25.502	50.849	6.598	
	4	36.028	70.560	3.759	
MDL2	8	1	1.712	3.6957	29.281
		2	2.408	5.938	25.143
		3	2.665	8.624	21.946
		4	2.867	11.758	19.274
	7	1	3.435	7.116	23.560
		2	5.332	12.548	18.643
		3	6.110	19.099	15.026
		4	6.824	26.232	12.356
MDL3	8	1	1.712	3.697	29.285
		2	2.407	5.957	25.170
		3	2.665	8.652	21.974
		4	2.866	11.797	19.303
	7	1	3.433	7.119	23.563
		2	5.330	12.576	18.663
		3	6.108	19.161	15.026
		4	6.822	26.330	12.356

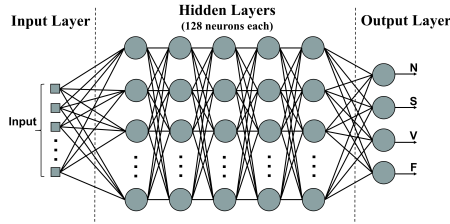


Fig. 2. Proposed artificial neural network for event driven classification

44 records of the MIT-BIH Arrhythmia database (excluding the 4 paced records as recommended by AAMI standards [14]). As observed in [8], we found that for $M > 8$, the LC-ADC produces more samples than a standard uniform ADC for MDL2 and MDL3. Therefore, these models have only been tested at $M=7$ and 8 . For higher resolutions in MDL2 and MDL3, k must be increased beyond 4LSB to attain compression, which is not tested in this study.

Both MDL2 and MDL3, pass the "good" signal criteria defined in (6) at $M=8$ ($k=1-3$ LSB) or $M=7$ ($k=1$ LSB). Whereas, MDL1 passes the criteria with $M=11$ ($k=1-4$ LSB), $M=10$ ($k=1-3$ LSB) or $M=9$ ($k=1$ LSB).

III. ARRHYTHMIA CLASSIFICATION USING EVENT-DRIVEN ECG DATA

A. Artificial Neural Network Based Classifier

We designed an ANN with 5 hidden layers consisting of 128 neurons each as shown in Fig. 2 for patient-specific (PS) arrhythmia classification using Python. The size of the input

TABLE II
ARRHYTHMIA CLASSIFICATION USING EVENT-DRIVEN DATA

Model	M	k	FS	ACC(%)	SE(%)	+P(%)	FPR(%)
MDL0	-	-	302	98.41	91.32	92.24	1.82
MDL1	11	1	76	97.95	86.89	88.92	3.68
		2	76	97.91	87.71	89.96	3.10
		3	76	97.91	86.88	88.12	3.75
		4	76	97.60	83.78	88.42	4.48
	10	1	76	98.21	87.68	90.10	3.48
		2	76	97.83	84.35	89.19	4.14
		3	62	97.05	81.75	85.79	5.60
		4	34	96.64	82.99	83.20	5.77
	9	1	62	97.45	83.56	86.81	4.42
		2	34	96.64	80.94	82.84	5.37
		3	26	95.83	78.42	80.12	7.53
		4	26	95.80	78.04	79.23	7.69
	8	1	34	96.63	81.57	83.62	5.98
		2	20	95.20	76.06	77.28	8.46
		3	16	94.25	73.87	77.83	9.78
		4	12	92.95	70.28	72.01	10.40
MDL2	8	1	172	96.66	84.24	82.72	5.22
		2	152	95.77	80.88	80.36	6.87
		3	142	94.93	78.27	77.70	7.36
		4	132	94.60	78.81	75.62	8.08
	7	1	86	95.96	79.46	80.18	6.47
		2	82	94.91	77.36	77.11	7.94
		3	78	94.60	76.03	75.14	8.43
		4	72	94.57	74.10	73.24	9.40
MDL3	8	1	152	96.52	82.54	82.86	5.06
		2	136	95.04	80.71	78.34	6.87
		3	136	94.42	78.10	78.38	8.49
		4	132	94.21	77.93	74.70	8.39
	7	1	102	96.47	81.24	81.91	5.79
		2	78	94.78	76.11	76.50	8.14
		3	74	94.38	74.94	74.50	8.84
		4	64	93.03	73.18	71.21	9.89

layer varies for each model based on the number of features. The output layer classifies each beat into one of the following 4 categories: normal beat (N), supraventricular ectopic beats (S), ventricular ectopic beats (V), and fusion beats (F). The prominent morphological features of an ECG beat is contained in an 800ms window [11] centered around the QRS peak as shown in Fig. 1(b). For every LC-ADC model, samples that lie within 400ms on each side of the QRS peak are used as the first feature vector. The pre-RR and post-RR intervals for each beat are used as the last two features. The conditional grouping scheme defined in [11] is applied in this design, which excludes 14 records from evaluation, in addition to the 4 paced records. To handle the inherent class imbalance in the dataset, the output layer uses a weighted cross-entropy loss, which balances the loss based on the effective number

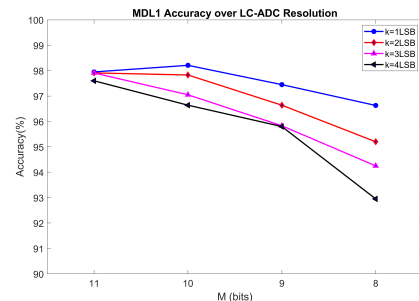


Fig. 3. MDL1 Accuracy at different LC-ADC resolutions (M) & varying k.

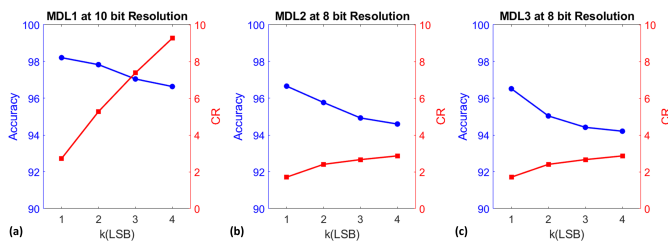


Fig. 4. The accuracy and compression performance of different LC-ADC models over varying k . (a) MDL1 at $M=10$ bit, (b) MDL2 at $M=8$ bit, and (c) MDL3 at $M=8$ bit

of samples in each class [15].

B. Classification Performance

For all records, there is 70:30 data split for training and testing the ANN classifier. First, the training data from all records are used for global training of the ANN. It is then retrained on the PS training data and evaluated on every individual subject's data. For a fair comparison, we have also tested the classifier on uniformly-sampled data using the MIT-BIH database as is, named MDL0 from here on. The MIT-BIH database is sampled at 360Hz using an 11-bit ADC. The accuracy (ACC), sensitivity (SE), positive predictivity (+P), and false positive rate (FPR) are used to assess the classifier performance as in [11]. The average results overall beat types (N, S, V, F) using each LC-ADC model and MDL0 are shown in Table II. Here, FS refers to the input feature size of the neural network. It can be observed that MDL1 at 11 bit, $k=2$ LSB gives the closest performance to MDL0, with almost 75% fewer features. Fig. 3 shows the MDL1 average accuracy over all beats as M is varied for different k values. At $M=11$, all LC-ADC resolutions have similar accuracy. However, as M is decreased, the difference between accuracy at various k values increases with the lowest accuracy at $k=4$ LSB for $M=8$. Fig. 4 shows a comparison of CR with accuracy over varying k . In Fig. 4(a) MDL1 at $M=10$ has the best accuracy over the range of k , also achieving higher CR as k is increased. In Fig.4(b)-(c) for MDL2 and MDL3 at $M=8$, the accuracy range, and CR are a little lower in comparison with MDL1.

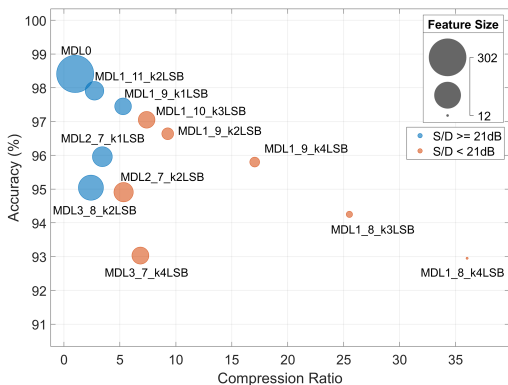


Fig. 5. Accuracy Vs CR Vs Complexity for different sampling schemes

C. Finding the Ideal Accuracy-Complexity Tradeoff

Input feature size (FS) of the ANN model is a good parameter to estimate the computational and hardware complexity of the classifier. MDL0, which uses uniformly sampled data from a standard 360Hz 11-bit ADC, requires an FS of 302. For event-driven classification to be effective, all other models should attain a significantly lower FS with similar or slightly lower accuracy. To understand the ideal trade-off between these parameters (i.e., accuracy, S/D, CR, and the complexity of the ANN), we picked the best models from Table II. Fig. 5 shows a bubble chart of the top 12 models. Among the LC-ADC models whose signal can be reconstructed with good quality (as in (6)), MDL1 at $M=9$, $k=1$ LSB gives the highest CR (5.277) and lowest complexity (FS=62) with over 97% accuracy. MDL1 at $M=11$, $k=2$ LSB offers the next best tradeoff with FS=76 and 98% accuracy at CR (2.722). If reconstructing the original ECG signal is considered not essential, then MDL1 at $M=8$, $k=4$ LSB has the highest compression (36.03) and lowest complexity (FS=12) with an acceptable 93% accuracy. Similarly, MDL1, $M=8$, $k=3$ LSB offers slightly lower compression (25.5) with over 94% accuracy and FS=16.

D. Open-Source Event-Driven ECG Dataset

To the best of our knowledge, there is no publicly available non-uniformly sampled ECG dataset with arrhythmia labels for researching event-driven data processing and classification. Therefore, in this study, we used the popular MIT-BIH arrhythmia database and three different LC-ADC models to derive a non-uniformly sampled event-driven ECG dataset. The dataset is generated using a MATLAB script and can generate LC-ADC data at different bit resolutions and k values. The event-driven ECG dataset with arrhythmia annotations and corresponding MATLAB scripts is open-sourced in authors' website¹ to enable further research on the topic.

IV. CONCLUSIONS

Detailed performance analysis of three different LC-ADC designs was presented. It was observed that increasing the quantization step size in an LC-ADC results in higher compression at the cost of lower signal-to-distortion ratios. Therefore, the signal quality of event-driven signals must always be assessed to ensure that the ECG beats are of clinically acceptable quality. Further, a 5-layer ANN architecture with class-balanced cross-entropy loss was proposed for the classification of event-driven ECG data. For the event-driven ECG arrhythmia classification, we achieved comparable classification accuracy to that of uniformly-sampled signals, whilst significantly decreasing the ANN complexity. Finally, we developed an open-source event-driven ECG database with arrhythmia annotations for enabling further research on non-uniformly sampled data processing.

¹The open-source event-driven ECG dataset is available at <https://github.com/jedaiproject/Open-Source-Event-Driven-ECG-Dataset>

REFERENCES

- [1] D. L. T. Wong, J. Yu, Y. Li, C. J. Deepu, D. H. Ngo, C. Zhou, S. R. Singh, A. Koh, R. Hong, B. Veeravalli, M. Motani, K. C. Chua, Y. Lian, and C. Heng, "An integrated wearable wireless vital signs biosensor for continuous inpatient monitoring," *IEEE Sensors Journal*, vol. 20, no. 1, pp. 448–462, 2020.
- [2] F. Chen *et al.*, "Design and Analysis of a Hardware-Efficient Compressed Sensing Architecture for Data Compression in Wireless Sensors," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 3, pp. 744–756, March 2012.
- [3] J. Li, A. Ashraf, B. Cardiff, R. C. Panicker, Y. Lian, and D. John, "Low power optimisations for iot wearable sensors based on evaluation of nine qrs detection algorithms," *IEEE Open Journal of Circuits and Systems*, vol. 1, pp. 115–123, 2020.
- [4] C. J. Deepu, X. Y. Xu, D. L. T. Wong, C. H. Heng, and Y. Lian, "A 2.3 μ w ecg-on-chip for wireless wearable sensors," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 65, no. 10, pp. 1385–1389, 2018.
- [5] J. Van Assche and G. Gielen, "Power Efficiency Comparison of Event-Driven and Fixed-Rate Signal Conversion and Compression for Biomedical Applications," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 14, no. 4, pp. 746–756, 2020.
- [6] Y. Hou *et al.*, "A 61-nW level-crossing ADC with adaptive sampling for biomedical applications," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 66, no. 1, pp. 56–60, 2018.
- [7] C. Weltin-Wu and Y. Tsvividis, "An event-driven clockless level-crossing ADC with signal-dependent adaptive resolution," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 9, pp. 2180–2190, 2013.
- [8] M. Tlili, M. Ben-Romdhane, A. Maalej, F. Rivet, D. Dallet, and C. Rebai, "Level-crossing ADC design and evaluation methodology for normal and pathological electrocardiogram signals measurement," *Measurement*, vol. 124, pp. 413–425, 2018.
- [9] X. Zhang and Y. Lian, "A 300-mV 220-nW Event-Driven ADC With Real-Time QRS Detection for Wearable ECG Sensors," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 8, no. 6, pp. 834–843, 2014.
- [10] N. Ravanshad *et al.*, "A Level-Crossing Based QRS-Detection Algorithm for Wearable ECG Sensors," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 1, pp. 183–192, 2014.
- [11] Y. Zhao, Z. Shang, and Y. Lian, "A 13.34 μ W Event-Driven Patient-Specific ANN Cardiac Arrhythmia Classifier for Wearable ECG Sensors," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 14, no. 2, pp. 186–197, 2019.
- [12] A. L. Goldberger *et al.*, "Physiobank PhysioToolkit and PhysioNet components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.
- [13] Y. Zigel *et al.*, "The weighted diagnostic distortion (WDD) measure for ECG signal compression," *IEEE transactions on biomedical engineering*, vol. 47, no. 11, pp. 1422–1430, 2000.
- [14] A.-A. EC57, A. for the Advancement of Medical Instrumentation *et al.*, "Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms," *Association for the Advancement of Medical Instrumentation, Arlington, VA*, 1998.
- [15] Y. Cui, M. Jia, T.-Y. Lin, Y. Song, and S. Belongie, "Class-balanced loss based on effective number of samples," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 9268–9277.